

基本句・内容語の説明書

河原 大輔

黒橋 禎夫

2008 年 8 月 5 日

1 基本句

基本句は格・省略解析などにおける処理の基本単位である。形態素を以下のように分類する。基本句は、内容語 1 つを核とする形態素列である。ただし、英語の句の場合のみ、1 基本句内に内容語が複数存在する。

接頭	自立	接尾	
内容語 前、元	内容語 並み、直前、途中 直後、予定	内容語 者、率、増 前、後、以前 以降、以来、中 勝、泊、角形、	1基本句 = 1内容語 例外 (1基本句に複数の 内容語) GREAT VIEW
		準内容語 個、cm、県 性、化、的だ さ、み 月、助数辞	
none	none	none	
お、ご	自身、全体、ともども	さん、君、ら、側、等	

2 内容語・準内容語

<内容語>や<準内容語>とは、意味属性を取得したり、キーワードとして用いるべき形態素である。<準内容語>は、基本句として独立しないが、何かしらの意味を保持する接尾辞である。

※ 従来は<意味有>という feature であったが、<内容語>または<準内容語>に改名された。

3 独立語、内容語の観点からの接頭・接尾辞の分類

以下のように、接頭辞は2つ、接尾辞は3つに分かれる。

- 接頭辞

- － <内容語>
- － それ以外

- 接尾辞

- － <内容語>
- － <準内容語>
- － それ以外

以下に、それぞれに属する形態素リストを示す。

- 接頭辞

- － <内容語>
 - * 名詞接頭辞
再 準 新 旧 超 高 低 猛 全 現 前 元 故 好 長 短 諸 被 真 真 っ 真 ん 大 中
小 両 本 同 各 今 来 最 他 初 副 名 昨 毎 翌 実 主 総 プレ
- － 非内容語
 - * 名詞接頭辞
御 お み す 第 約 ま る ひ と
 - * 動詞接頭辞、イ形容詞接頭辞、ナ形容詞接頭辞のすべて

- 接尾辞

- － <内容語>
 - * 名詞性名詞助数辞
社 会 条 項 単 位 種 類 連 続 余 り 戦 勝 敗 泊 角 形
 - * 名詞性特殊接尾辞
増 減 暮 れ 末 以 前 以 後 以 来 以 降
 - * 名詞性名詞接尾辞
前 中 後 前 後 率 通 り 型 形 用 製 家 者 数 代 費 学 向 き 向 け 入 り 明 け
作 り 造 り づ く り ず み 無 し な し 行 き 宛 あ て ど こ ろ
 - * 名詞性述語接尾辞
方 手 放 題
- － <準内容語>
 - * 名詞性名詞助数辞のうち<内容語>でないもの
 - * 名詞性特殊接尾辞

都道府県郡市区州省

* 名詞性名詞接尾辞

性化以外

* 名詞性述語接尾辞

さ味み

* 形容詞性名詞接尾辞

的だ がちだ 気味だ

* 形容詞性述語接尾辞

目だ めだ 気だ げだ

－ 非内容語

* 名詞性特殊接尾辞

超 以内 限り 過ぎ 半 強 弱 付

* 名詞性名詞接尾辞

さん 君 様 殿 氏 ちゃん 以上 以下 未満 近く くらい ぐらい 等 程 間 程度
頃 ごろ ずつ 当たり 辺り おき 毎 振り 振 ぶり っ ぶり 目 だらけ さ な か
かたがた 自体 達 ら 同士 ども 方 勢 っ 子 調 上 下 そのもの が かり
足らず たらず がらみ ぐるみ ならでは ぞろい がえ 側 ずくめ 尽くし
つくし ずく ごとき 如き 内 その他 級 づたい 別

* 名詞性述語接尾辞

たて 振り 振 ぶり っ ぶり 様 が い っ こ

* 形容詞性名詞接尾辞

く さい っ ぽい い

* 形容詞性述語接尾辞

ない たい やすい よい にくい がたい 難い づらい がちだ そうだ

* 動詞性接尾辞のすべて

4 問題

- 一つの基本句に複数の<内容語>(または<準内容語>)がある場合に検索で問題発生

－ <内容語>+<準内容語>

例) 3 個, 3 冊, 東京都, 可能性, 国際化, 長さ, 積極的だ

→ 検索側で対処